白菜 *CesA* 基因家族鉴定及表达模式分析

马宇辰[1]，赵玉梅[1, 2]，黄丹霖[1]，张梦晴[1]，吴晓宇[1, 2]，王洁[1, 2]，
陈雨[1]，黄家保[1, 2]，段巧红[1, 2*]
（1. 山东农业大学园艺科学与工程学院，泰安 271000；2. 作物生物学国家重点实验室，
泰安 271000）

摘要：为探究 *CesA* 基因家族在白菜生长发育及纤维素合成过程中的作用机制，该研究通过生物信息学的方法，以白菜的全基因组序列为研究区域，进行了理化特征、基因结构、进化特征、保守基序及结构域、顺式作用元件和组织表达等鉴定分析。结果表明:（1）白菜基因组中鉴定出 16 个编码纤维素合成酶亚基的 *CesA* 基因，该家族成员所编码蛋白的理论等电点介于 4.76~9.12，相对分子量 17.76~122.67 kD，长度 153~1 089 aa；（2）其中 15 个基因不均匀地分布于白菜的 7 条染色体上，*Bra*036008 定位于 scaffold 上；（3）大部分成员包含 4-14 个外显子，1-11 个保守基序；（4）序列比对显示该家族具有保守的 DDD-QXXRW 保守功能域；（5）该家族编码蛋白主要分布在质膜上，二级结构以无规则卷曲与α-螺旋为主，多数成员都含有 CesA 蛋白典型的 N 端、C 端和跨膜区；（6）*CesA* 基因在茎中表达量相对较高，其中 *Bra*011865、*Bra*023952 和 *Bra*029874 在茎、叶、花中显著表达。该研究利用生物信息学方法对白菜 *CesA* 基因家族进行了全基因组鉴定，为后续深入研究 *CesA* 基因功能奠定了基础，也为白菜生长发育研究奠定基础。
关键词：纤维素合成酶，白菜，*CesA*，基因家族，基因表达

Identification and expression analysis of *CesA* gene family in *Brassica rapa*

MA Yuchen[1], ZHAO Yumei[1, 2], HUANG Danlin[1], ZHANG Mengqing[1], WU Xiaoyu[1, 2],
WANG Jie[1, 2], Chen Yu[1], HUANG Jiabao[1, 2], DUAN Qiaohong[1, 2*]
（1. *College of Horticultural Science and Engineering, Shandong Agricultural University*, Tai'an271000, Shandong, China；2. *State Key Laboratory of Crop Biology*, Tai'an 271000, Shandong, China）

Abstract Cellulose is the main component of plant cell walls. It is involved not only in cell morphology and development, but also in various cellular signal transduction pathways, thus

基金项目：山东省自然科学基金重点项目（ZR2020KC017）[Supported by the Natural science fund project in Shandong Province (ZR2020KC017)]。
作者简介：马宇辰(2000-)，女，本科生，园艺科学与工程，（E-mail）2586922861@qq.com。
***通信作者**：段巧红，博士，教授，研究方向为远源杂交中生殖隔离的形成机理，（E-mail）duanqh@sdau.edu.cn。

affecting plant growth and development. Cellulose synthase is the main enzyme of cellulose synthesis in plants. In order to explore the role of *CesA* gene family in cellulose synthesis and development of *Brassica rapa*, we identified the members of *CesA* genes in *B. rapa* genome via bioinformatic method. The results were as follows：(1) We identified 16 *BrCesA* genes from 10 *B. rapa* chromasomes with pl range from 4.76-9.12, molecular weight range from 17.76-122.67 kD and amino acid length range from 153-1 089 aa. (2) With an expectation of *Bra036008*, which located in scaffold, the rest of 15 *BrCesA* genes unevenly distributed in 7 chromosomes. (3) Most of *CesA* contained 4-14 exons and 1-11 conserved motifs. Sequence alignment showed that all *CesAs* contained DDD-QXXRW domain. (4) Subcellular location analysis presents a typical N terminal, C terminal and transmembrane structure and the protein showed plasma membrane location. (5) Tissue specific analysis by qRT-PCR showed relative highly expression in stem of *B. rapa* seedlings. (6) When seedling treated with 150 mM·L-1 NaCl, the expression of *Bra01069* significantly increased after 6 h. At present, although *B. rapa* reference genome sequencing research has advanced to the new version v3.0, the *CesA* gene family has not been thoroughly identified and analyzed in *B. rapa*, and the functions of many genes have not yet been clarified. In this study, bioinformatics methods were used to take the whole genome sequence of *B. rapa* as the research field, and the whole genome identification of the *CesA* gene family of *B. rapa* can not only lay the foundation for the subsequent in-depth study of the function of *CesA* gene, but also lay the foundation for the growth and development of *B. rapa*.

Keywords　Cellulose synthase，*Brassica rapa*，*CesA*，gene family，gene expression

纤维素（cellulose）是植物细胞壁的主要成分（姚敦义等，1988），在细胞生长、分化、信号转导等过程中起到重要作用(Xie et al., 2021)。纤维素合酶（Cellulose synthases A, CesA）是合成纤维素的酶（Doblin et al., 2002），本质上属于糖基转移酶，最初在革兰氏阴性细菌醋杆菌中被鉴定（Wong et al., 1990），植物纤维素合酶是由 36 个单体组成的玫瑰状复合体，其单体主要由纤维素合成酶基因家族成员编码。高等植物纤维素合成酶复合体中，有不少于三种 CesA 亚型参与（Doblin et al., 2002），CesA 亚基彼此相互作用形成高级复合物。

迄今为止，*CesA* 基因家族已经在模式植物拟南芥(Pear et al., 1996)、水稻(Huang et al., 2015)、烟草（Xu et al., 2017）中鉴定出来，其中拟南芥 *CesA* 基因的研究最为深入(Richmond, 2000)。研究表明，拟南芥基因组拥有 10 个 *CesA* 基因(*CesA1-10*)，它们编码与细菌纤维素合成酶同源的蛋白(Pear et al., 1996)，其次生细胞壁的形成需要三种类型的 CesA 亚基(CesA4/不规则 XYLEM5 [IRX5]，CesA7/IRX3 和 CesA8/IRX1)(Hernández-Blanco et al., 2007)，*CesA*3 基因突变使纤维素合成水平降低，通过茉莉酸和乙烯等信号通路激活木质素合成和防御反应（Caño-Delgado et al., 2003）；在水稻基因组中共鉴定到 11 个 *OsCesA* 基因，水稻中的赤霉素信号转导可以通过促进纤维素的合成，进而影响水稻植株节间的发育(Huang et al., 2015)；此外，烟草基因组中鉴定出 21 个 *NtCesA* 基因，*NtCesA*7、*NtCesA*9、*NtCesA*11、*NtCesA*16、*NtCesA*18 等基因主要参与烟草次生细胞壁纤维素的合成（Xu et al., 2017）。除模式植物外，玉米(Appenzeller et al., 2004)、棉花（Meng et al., 2012）、巨龙竹（Wang et al., 2021）、紫花苜蓿 (Liu et al., 2018）、罗布麻(Xie et al., 2021)等多种非模式植物中也鉴定出 *CesA* 基因。从玉米中鉴定出 12 个不同的 *ZmCesA* 基因，在初级或次级细胞壁合成的独特细胞类型中表达(Doblin et al., 2002)。棉花中鉴定出 18 个 *CesA* 基因（Meng et al., 2012）；巨龙竹中鉴定出 45 个 *CesA* 基因（Wang et al., 2021）。紫花苜蓿盛花期 *CesA* 基因表达水平显著增长，纤维素含量大幅增加（Liu et al., 2018）。罗布麻 *CesA* 基因家族包含的大量光响应、胁迫响应、

激素响应等元件，在植株生长发育与抗逆防御方面发挥着作用(Xie et al., 2021)。

白菜（*B. rapa*）属十字花科（*Cruciferae*）芸薹属（*Brassica*）二年生草本植物，原产于我国，栽培历史悠久，年种植面积 180 万 hm²左右(张凤兰等，2021)，是人们日常餐桌上必不可少的蔬菜。2011 年，白菜全基因组测序工作完成（Wang et al., 2011），2017 年获得了升级的 v2.5 版本（Cai et al., 2017），2018 年获得了白菜参考基因组 v3.0 新版本（Zhang et al., 2018），为后续开展白菜基因研究奠定了坚实的基础（Chen et al., 2020）。目前 *CesA* 基因家族在白菜中还未进行深入鉴定与分析。本研究以白菜的全基因组序列为研究区域，依托在线软件与 qRT-PCR 分析，采用生物信息学方法，通过对 *CesA* 基因家族的结构、进化特征、组织表达等进行系统鉴定，拟探讨以下问题：（1）白菜 *CesA* 基因家族的遗传进化规律；（2）该家族成员的理化性质；（3）该家族成员的基因结构、蛋白质二级结构及保守结构域；（4）该家族成员在各组织中的表达模式以及在盐胁迫下的表达模式。

1 材料与方法

1.1 基因家族成员鉴定及理化性质分析

根据拟南芥已经鉴定的 10 个 *CesA* 基因序列，通过 BLAST 在白菜基因组（*B. rapa* genome v1.5）中搜索候选的白菜 *CesA* 家族成员，并通过综合分析确定候选基因。使用 TBtools 软件(Chen et al., 2020)结合 Expasy(Wang et al., 2012)（https://www.expasy.org/）分析白菜 *CesA* 基因家族的蛋白质分子量(MW)、等电点(pI)等理化性质。

1.2 基因结构及保守结构域分析

利用 TBTools 软件绘制白菜 *CesA* 基因结构图；利用 MEME（Kumar et al., 1994）（http://meme-suit.org/tools/meme）在线预测 CesA 蛋白质基序。TBtools 软件(https://github.com/) 绘制可视化 MEME 结构，并进行基因结构分析。

1.3 基因在染色体上的分布及共线性关系分析

通过白菜基因组数据库对鉴定获得的 *CesA* 基因家族成员进行染色体分析，利用 TBTools 软件作图进行分析。使用 BLAST 比对，合并、建库、蛋白对比；运行 MCScanX(Bolser et al., 2016)，进行下游分析及可视化作图，从而进行共线性分析，得到共线性图。

1.4 亚细胞定位预测和蛋白质二级与三级结构分析

使用 WoLF PSORT（http://www.genscript.com/wolf-psort.html) 进行亚细胞定位 PRABI(https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html)进行蛋白二级结构分析。利用 Phyre2 (Kelley et al., 2015)（http://www.sbg.bio.ic.ac.uk/ phyre2/ html/page.cgi?id=index）对白菜 CesA 蛋白进行三级结构预测。

1.5 白菜 *CesA* 基因家族成员蛋白跨膜螺旋与三级结构预测

利用 TMHMM Server v2.0（http://www.cbs.dtu.dk/services/TMHMM/）对白菜 CesA 蛋白跨膜螺旋进行预测。

1.6C 白菜 *CesA* 基因表达分析

以“蛋黄白”白菜品种为实验材料。提取 30 d 苗龄的根、茎、叶和 80 d 苗龄的花和果荚的 RNA，进行组织特异性表达分析。挑选生长状况相近的白菜幼苗浸泡于用 Hoagland 营养液配制的 150 mmol • L^{-1} NaCl 溶液中，12 h 后进行材料的 RNA 提取，并以正常水培幼苗为对照，进行盐胁迫下基因表达分析。

使用 qPrimerDB-qPCR Primer Database（https://biodb.swu.edu.cn/qprimerdb/）(Li et al., 2017)设计 qRT-PCR 引物，以 *BraActin 2* 为内参基因（表 1）。使用 RNA 提取试剂盒提取样品总 RNA，使用 HiScript ll QRT SuperMix for qPCR 试剂盒反转录得到 cDNA。使用 ChamQ Universal SYBR qPCR Master Mix 试剂盒进行定量实验。反应体系 20 μL:2 μL cDNA，上、下游引物各 0.4 μL, SYBR 10 μL，用灭菌 dd H_2O 补至 20 μL。PCR 反应程序：95 ℃预变性

30 s,95℃变性 10 s，60℃退火 22 s，进行 40 个循环。基因的相对表达量以 $2^{-\Delta\Delta Ct}$ 法计算。所得数据通过 TBtools 进行可视化分析。

表 1 白菜 *CesA* 基因家族表达分析的实时荧光定量 PCR 引物

Table 1 qRT-PCR primers for expression analysis of *CesA* gene family in *Brassica rapa*

基因编号 Gene ID	引物序列（5'—3'） Primer sequence	
Bra023952	F : AAAGGTGGAGAAATTGAAGGGA	R : AATAATCACAACCCGGTAAGGT
Bra011345	F : GACATGATTCTCAGCCAATTCC	R : CTGATGTAGTTCGCACAGATTG
Bra011865	F : ATCGATACTCTAATCGCAACGT	R : ACAACCTGTCCCGACATATATC
Bra010691	F : CGCCTCGTCTAGTATTAACTGT	R : CCAACGTATATCGGTCCTTGTA
Bra029874	F : TTAGCTGTGGATATGAGGACAC	R : CAGTCTCCCTCCATAACCATAC
Bra005845	F : CGGAGCTCTACTTGTTCAAATG	R : GGTATCCACTGTTGATAGCGTA
Bra028768	F : CATGTCATTAGCTGTGGTTACG	R : GCATTTTGAACCCAGTCAGAAT
Bra033714	F : GCAAGCTCAAGATTCTAGAACG	R : AACCGCTGGGATAGTACAATAG
Bra036008	F : CGACAACGGAAGCAATTTTCTA	R : GACATCCCAAATCTCTTCTCGA
Bra006036	F : CTTGGCCTGGTAATAATGTTCG	R : TTTGCTGTTGTTGATGTAGTGG
Bra037793	F : TAATAGCACAAGAGATCACCCC	R : GAACCCCTGAGACTCGTATAAG
Bra024324	F : TTTGTTCTGATTAATGCCGACG	R : TCTCTGCAGATTTTACACGTCT
Bra031904	F : ATATGAAGGGTTTGGATGGGTT	R : ATAGGAAACACCATTTAGGCCA
Bra004814	F : GAGGACGAGGATGATAATGAGG	R : ATACACTAACCCAACATCCCAG
Bra006407	F : CAAGCAAGATCAATCCATACCG	R : AGAGACGGTCTAGATAGGTCTC
Bra012578	F : ACAGTCTCCACAATAGACAGTG	R : GGATAGGTATCACAACCGAGAG
BraActin 2	F : CGGTGTCATGGTTGGGAGA	R : CGTGCTCGATGGGGTACTTC

2 结果与分析

2.1 白菜 *CesA* 基因家族成员的鉴定与理化性质分析

经筛选鉴定，得到 16 个白菜 *CesA* 基因家族成员（表 2），分析结果表明白菜 *CesA* 家族蛋白等电点介于 4.76~9.12 之间，酸性蛋白与碱性蛋白的数量相近，分别为 7 个与 9 个。分子量介于 17.76 kD~122.67 kD 之间。*Bra*011865 的氨基酸数目最多，*Bra*004814、*Bra*010691 蛋白的氨基酸数目显著少于其他成员。

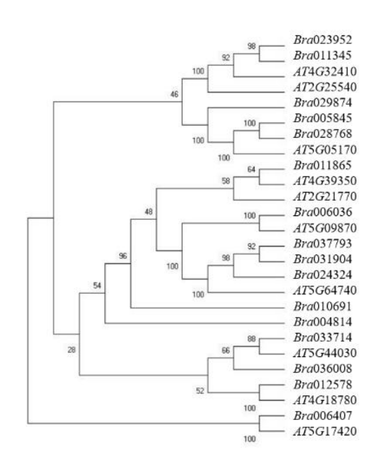
表 2 白菜 *CesA* 基因家族信息

Table 2 Information of *CesA* gene family in *Brassica rapa*

基因编号 Gene ID	染色体 Chromosome	等电点 Isoelectric point	分子量(kD) Molecular weight (kD)	起始位置(bp) Start location (bp)	终止位置(bp) End location (bp)	蛋白长度 Protein length (aa)
Bra023952	A03	6.34	122.17	28 621 459	28 626 059	1 083
Bra011345	A01	6.51	121.85	2 640 630	2 645 508	1 080
Bra011865	A01	7.04	121.03	150 169	154 808	1 074
Bra010691	A08	6.98	31.61	16 010 300	16 016 114	287
Bra029874	A05	8.39	118.03	21 907 336	21 912 132	1 046
Bra005845	A03	7.67	118.62	899 408	903 741	1 055
Bra028768	A02	7.43	119.79	2 136 518	2 141 039	1 066
Bra033714	A06	8.43	119.81	25 466 123	25 470 303	1 052
Bra036008	scaffold	9.12	20.92	492 903	493 889	1 089
Bra006036	A03	6.65	121.40	1 703 119	1 708 289	1 076
Bra037793	A09	7.65	122.18	3 605 709	3 610 867	1 083
Bra024324	A06	7.24	122.67	14 970 360	14 975 283	1 082
Bra031904	A02	7.81	121.49	27 187 433	27 192 926	1 079
Bra004814	A05	4.76	17.76	1 926 040	1 927 962	153
Bra006407	A03	6.18	116.43	3 382 190	3 386 798	1 034
Bra012578	A03	6.69	111.40	23 439 101	23 442 994	985

2.2 白菜 *CesA* 家族基因系统进化分析

许多物种中都有 *CesA* 家族同源基因的进化与表达（表 3）。为了解白菜 *CesA* 基因家族与拟南芥 *CesA* 基因家族之间的进化关系，利用 MEGA-X 软件，对两者的 *CesA* 基因家族成员进行了系统进化分析，结果表明白菜 *CesA* 基因家族与拟南芥 *CesA* 基因家族有相近的进化关系（图 1）。*CesA* 基因可以分为 4 个亚族，同时，共发现有 5 个白菜 *CesA*（*Bra*006407、*Bra*012578、*Bra*033714、*Bra*011865、*Bra*006036）与拟南芥 *CesA* 具有高度同源性，表明这五对基因具有高度的亲缘关系。通过对基因分化时间预测，在基因进化的角度推测以 *CesA*1 出现时间最早。



Bra. 白菜；**AT**. 拟南芥。下同。

Bra. *Brassica rapa;* **AT**. *Arabidopsis thaliana.* The same below.

图 1 白菜与拟南芥 CesA 蛋白家族系统进化树

Fig. 1 Distribution of conserved motifs of CesA protein family in *Brassica rapa* and *Arabidopsis thaliana*

表 3 *CesA* 家族同源基因在不同物种中的功能研究

Table 3 Studies on the function of *CesA* family homologous genes in different species

拟南芥基因 ID *Arabidopsis thaliana* ID	基因 ID Gene ID	物种 Species	功能 Function	参考文献 Reference
ATCesA2	*CrCESA2-2*	琯溪蜜柚 Guanxi pomelo	参与初生壁纤维素合成调控 Regulation of primary wall cellulose synthesis	Dai et al., 2021
ATCESA7、*ATCESA8*	*CrCESA4*、*CrCESA7-1*、*CrCESA8*	琯溪蜜柚 Guanxi pomelo	参与了汁胞粒化过程次生壁纤维素的积累 Accumulation of secondary wall cellulose during granulation of juice cells	Dai et al., 2021
ATCesA1、*ATCesA3*	*AvCesA2*、*AvCesA3*	罗布麻 *Apocynum venetum*	参与初生细胞壁纤维素合成基因 Primary cell wall cellulose synthesis gene	Xie et al., 2021
ATCesA7、*ATCesA4*、*ATCesA8*	*AvCesA5*、*AvCesA8*、*AvCesA10*	罗布麻 *Apocynum venetum*	参与次生细胞壁纤维素的合成 Primary cell wall cellulose synthesis	Xie et al., 2021
ATCesA6、*ATCesA1*	*De10040587*、*De10044972*、*De10083985*、*De10040239*	铁皮石斛 *Dendrobium candidum*	初生细胞壁的纤维素合成有关 Primary cell wall is involved in cellulose synthesis	Lan et al., 2019
ATCesA4、*ATCesA7*、*ATCesA8*	*De110044623*、*De10069041*	铁皮石斛 *Dendrobium candidum*	次生细胞壁合成相关 Secondary cell wall synthesis is related	Lan et al., 2019
ATCesA1、*ATCesA3*、*ATCesA6*	*MlCesA7*	南荻 *Miscanthus*	参与初级细胞壁的形成 Formation of primary cell wall	Li et al., 2021

		lutarioriparius		
ATCesA4、*ATCesA7*、*ATCesA8*	*MlCesA4*、*MlCesA9*	南荻 *Miscanthus lutarioriparius*	参与次级细胞壁的形成 Formation of secondary cell walls	Li et al., 2021
ATCesA8	*ClCesA1*	杉木 *Abies holophylla*	参与木材次生壁形成 Wood secondary walls form	Pang et al., 2015
ATCesA1	*ClCesA2*	杉木 *Abies holophylla*	参与细胞初生壁的形成 Formation of the primary wall of a cell	Pang et al., 2015
ATCesA1、*ATCesA2*、*ATCesA3*、*ATCesA5*、*ATCesA6*、	*LgCesA1*、*LgCesA3*、*LgCesA5 和 LgCesA6*	兴安落叶松 *Larix gmelinii*	与细胞初生壁的生长相关 Growth correlates with the primary wall of cells	Te et al., 2021
ATCesA4、*ATCesA7*、*ATCesA8*	*LgCesA4*、*LgCesA8*	兴安落叶松 *Larix gmelinii*	与细胞次生壁的生长相关 Growth correlates with cell secondary walls	Te et al., 2021
ATCesA3、*ATCesA9*	*F01_cb9914_c103031/f5p6/3907*、*F01__cb9914__c4357/f4p0/4200*	巨龙竹 *Dendrocalamus sinicus*	参与初生壁纤维素生物合成 Primary wall cellulose biosynthesis	Wang et al., 2021
ATCesA4、*ATCesA7*、*ATCesA8*	*NtCesA7*、*NtCesA9*、*NtCesA11*、*NtCesA16*、*NtCesA18*	烟草 *Nicotiana tabacum*	主要参与烟草次生细胞壁纤维素的合成 Synthesis of cellulose from secondary cell wall	Xu & Kong, 2017
ATCesA1、*ATCesA3*、*ATCesA6*	*NtCesA19*	烟草 *Nicotiana tabacum*	参与初生细胞壁纤维素合成 Primary cell wall cellulose synthesis	Xu & Kong, 2017
ATCesA1、*ATCesA3*、*ATCesA6*	*RpCesA2*	刺槐 *Robinia pseudoacacia*	参与初生细胞壁纤维素合成 Primary wall cellulose biosynthesis	Zhang et al., 2017
——	*DlCesA3-2*、*DlCesA6*、*DlCesA7-1*	龙眼 *Euphoria longan*	可能影响植物胚乳发育 Affect plant endosperm development	Zhu et al., 2020
——	*DlCesA7*	龙眼 *Euphoria longan*	可能参与龙眼体胚发育调控 Regulation of somatic embryo development	Zhu et al., 2020

注：拟南芥 ID 栏为物种基因家族对应同源基因；——表示无对应基因。

Note: ID column of *Arabidopsis thaliana* is the corresponding homologous gene of the species gene family; —— indicates no corresponding gene.

2.3 白菜 *CesA* 家族基因的结构和保守结构域分析

利用 TBtools 对 16 个白菜 *CesA* 基因家族成员进行基因结构分析，结果显示（图 2）。该家族成员大都包含 4~14 个外显子。其中 5 个基因含有 14 个 CDS 序列，3 个基因含有 13 个 CDS 序列，4 个基因含有 12 个 CDS 序列，*Bra*033714 含 11 个 CDS 序列，*Bra*010691 含 6 个 CDS 序列，*Bra*004814 含有 5 个 CDS 序列，*Bra*036008 含有 4 个 CDS 序列。表明部分家族基因结构间存在较大差异，具有丰富的多样性。

利用 MEME/MAST 程序进行蛋白保守结构域分析（图 3），共鉴定到 10 种 Motif，大部分成员含有相同的 Motif 组成。均含有一个保守的 Motif1；*Bra*011345 基因中含有 Motif1~10 及一个重复的 Motif6；*Bra*011865 基因中含有 Motif1~10 及一个重复的 Motif8，推测具有差异的基因在白菜生长周期中发挥不同于其他的作用。利用 DNAMAN 对白菜 *CesA* 进行氨基酸序列进比对显示（图 4），除了 *Bra*010691、*Bra*036008 和 *Bra*004814，其余家族基因都含有 CesA 蛋白典型 DDD-QXXRW 保守功能域，与拟南芥 CesA 蛋白具有相同的保守功能域(Delmer DP，1999)。

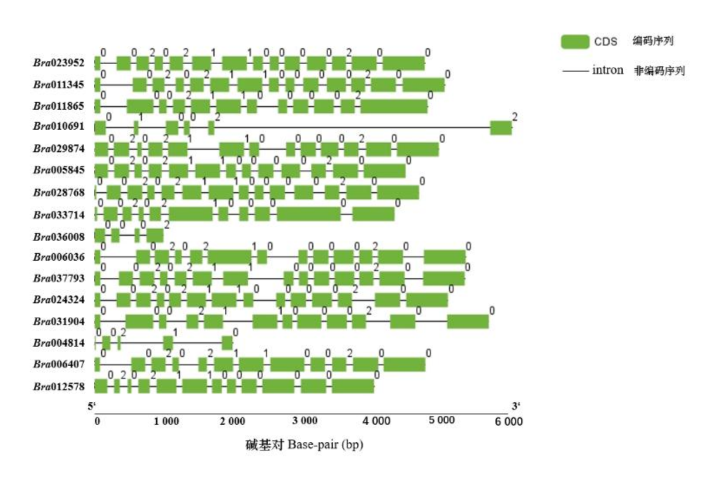


图 2 白菜 *CesA* 基因的结构

Fig. 2 Structure of *CesA* gene in *Brassica rapa*

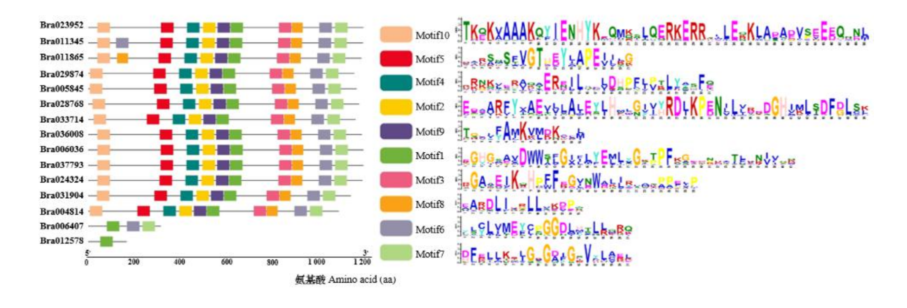
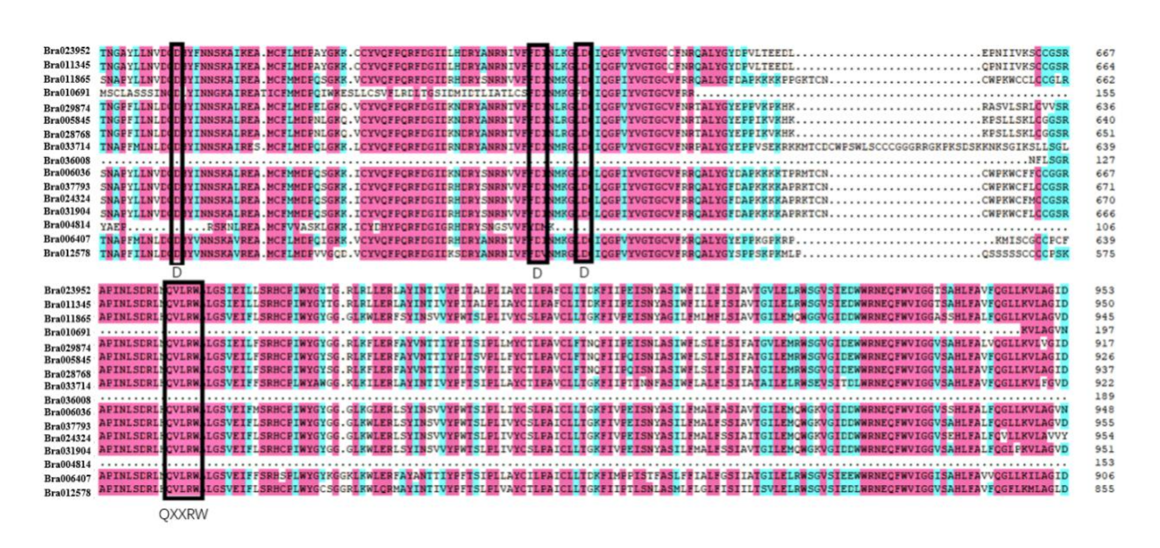


图 3 白菜 CesA 家族结构域分析

Fig. 3 Motif of CesAs family in *Brassica rapa*



框选内容为白菜 CesA 蛋白 D、DXD 保守功能域与 QXXRW 保守功能域。

Boxes are CesA protein D，DXD conserved functional domain and QXXRW conserved functional domain.

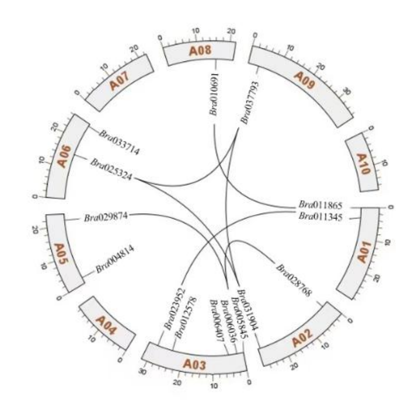
图 4 白菜 *CesA* 家族氨基酸序列比对

Fig. 4 Multiple secquence alignment of *CesA* family in *Brassica rapa*

2.4 白菜 *CesA* 家族染色体定位与共线性分析

利用 TBtools 绘制出白菜 *CesA* 家族基因的染色体分布图（图 5）。*CesA* 基因不均匀地分布于 A01、A02、A03、A05、A06、A08、A09 七条染色体上，*Bra*036008 定位于 scaffold 上。

利用 TBtools 对 16 个白菜 *CesA* 基因家族成员与拟南芥进行分析，共鉴定到 7 对共线性基因对（图 6）。大多数白菜 *CesA* 家族成员都可找到在拟南芥中的直系同源基因，仅 *Bra*010691 和 *Bra*004814 无与之对应的共线性的基因。

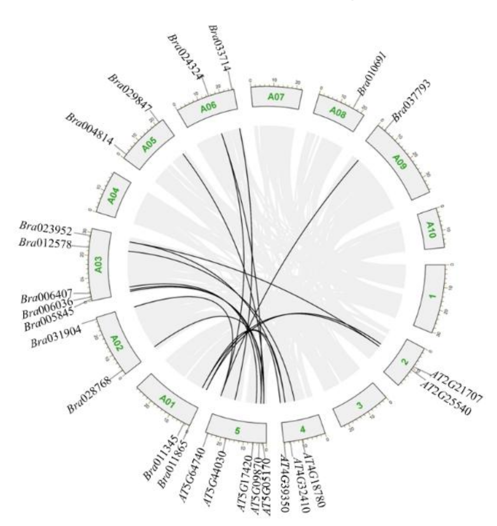


A01~A10 为白菜染色体。

A01~A10 are *Brassica rapa* chromosomes.

图 5 白菜 *CesA* 基因在染色体上的位置分布

Fig. 5 Chromosome localization of *CesA* gene in *Brassica rapa*



1~5 为拟南芥染色体；A01~A10 为白菜染色体。

1-5 are *Arabidopsis thaliana* chromosomes; A01-A10 are *Brassica rapa* chromosomes.

图 6 白菜与拟南芥间 *CesA* 家族共线性分析

Fig. 6 Synteny analysis of *CesA* family between *Brassica rapa* and *Arabidopsis thaliana*

2.5 亚细胞定位预测和蛋白质二级结构分析

基因的亚细胞定位有助于我们进一步了解其功能。利用 WoLF PSORT 在线网站进行亚细胞定位分析，发现绝大多数白菜 *CesA* 定位在质膜上，与纤维素合酶的作用机制相符。蛋白质二级结构是多肽链在空间的三维排列中的一个高级组织层次，了解 CesA 蛋白的结构有助于预测其分子功能。利用 PRABI 在线网站对白菜 CesA 蛋白进行分析，发现 *CesA* 基因家

族成员的蛋白质二级结构含有大量无规则卷曲、α-螺旋和延伸链结构及少量的β-转角（表 3）。

表 3 白菜 *CesA* 亚细胞定位预测与二级结构分析

Table 3 Prediction of subcellular location and secondary structure of *CesA* in *B. rapa*

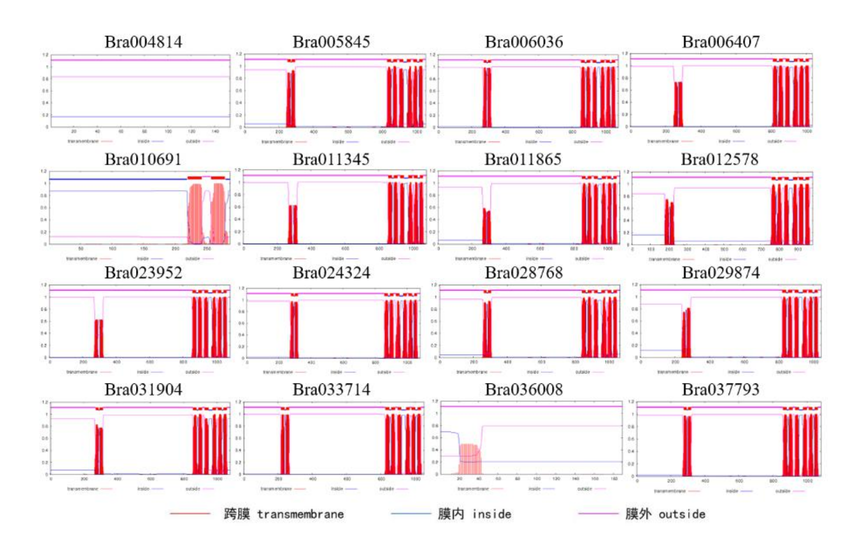
蛋白质 ID Protein ID	亚细胞结构定位 Sbcellular localization prediction							蛋白质二级结构分析 Protein secondary structure			
	质膜 Plas	叶绿体 Chlo	胞浆 Cyto	线粒体 Mito	液泡 Vacu	细胞核 Nucl	内质网 E.R.	α-螺旋 α-helix (%)	延伸链 Extended strand (%)	β-转角 β-turn (%)	不规则卷曲 Random coil (%)
Bra023952	11						2	31.86	15.51	4.89	47.74
Bra011345	12					1		32.13	16.67	4.63	46.57
Bra011865	12					1		33.71	15.74	5.59	44.97
Bra010691	7		1	1	2	1	1	35.54	24.04	9.06	31.36
Bra029874	9					2	2	32.7	14.63	4.49	48.18
Bra005845	10			1		2		33.27	14.6	4.36	47.77
Bra028768	11					2		33.49	14.45	4.69	47.37
Bra033714	11					2		35.08	14.54	5.23	45.15
Bra036008		1		4		8		28.04	17.46	8.47	46.03
Bra006036	13							31.78	15.24	4.28	48.7
Bra037793	13							31.95	15.14	4.89	48.01
Bra024324	13							32.75	14.37	4.2	48.68
Bra031904	13							32.25	14.02	3.73	50
Bra004814		8	4			1		16.99	26.14	10.46	46.41
Bra006407	11	1				1		34.53	14.81	3.59	47.06
Bra012578	12					2		32.79	14.42	4.77	48.02

注：关于亚细胞定位数值根据分选信号、氨基酸组成和功能域数，经 kNN 算法而得到；数值表示所占比例；“—”表示无数值

Note: Values of subcellular localization is obtained according to sorting signals，amino acid composition and functional motifs from KNN. The numerical value represents the proportion column; — represents the absence of a value.

2.6 白菜 *CesA* 家族成员蛋白跨膜结构域预测

利用 TMHMM 对白菜 *CesA* 家族成员蛋白的跨膜结构进行预测，结果显示（图 7）：除 *Bra*010691、*Bra*036008、*Bra*004814 蛋白外，其余家族成员蛋白在 C 端均含有 6 个完整的跨膜结构域，且每个组成跨膜结构域的氨基酸数量与位置均相似，据此推测其可能有相同的结构和功能，而 *Bra*010691、*Bra*036008、*Bra*004814 可能因为发生突变，相应功能发生变化，而导致跨膜结构缺失。仅有 *Bra*005845、*Bra*006036、*Bra*024324、*Bra*028768、*Bra*031904、*Bra*033714、*Bra*037793 蛋白在近 N 端均含有 2 个完整的跨膜结构域。



横坐标为蛋白序列长度（氨基酸数量）；纵坐标为概率。

The x-coordinate is the length of the protein sequence (the number of amino acids); the y-coordinate is the probability.

图 7 白菜 CesA 蛋白跨膜螺旋预测

Fig. 7 Prediction of CesA protein transmembrane helix of *Brassica rapa*

2.7 白菜 *CesA* 基因家族表达分析

通过 qRT-PCR 分析 *CesA* 在根、茎、叶、花、果荚中表达水平，使用 TBtools 将这些基因进行可视化分析（图 8）。结果表明 *CesA* 基因在根、茎、叶、花、果荚中均有不同程度的表达，且在茎部表达量较高。*Bra*011865、*Bra*023952 和 *Bra*029874 在茎、叶、花和果荚中表达量均显著高于在根中的表达水平，其中，*Bra*029874 的表达量最为显著，在茎、叶、花中表达量为在根中表达量的 6~7 倍。

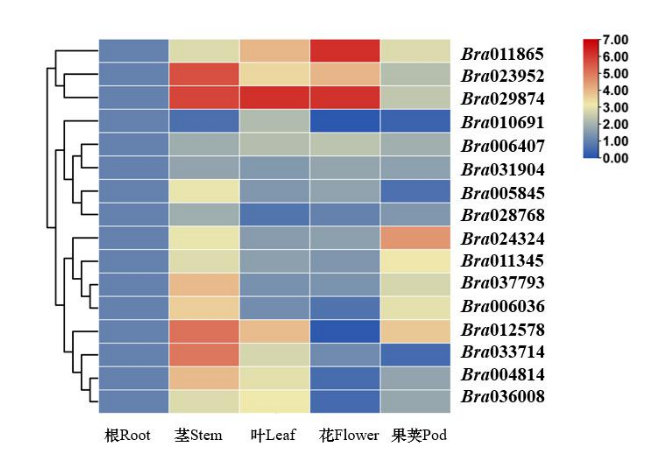


图 8 白菜不同组织中 *CesA* 基因的表达模式

Fig.8 Expression pattern of *CesA* genes in different tissues

通过 qRT-PCR 分析 *CesA* 在正常与 150 mM·L^{-1} NaCl 胁迫状态下的表达水平，结果表明大部分 *CesA* 基因对盐处理有一定程度的响应，其中 *Bra*011865 的 2 h 处理和 6 h 处理组，与 *Bra*010691 的 6 h 处理组响应最为显著，*Bra*024324、*Bra*028768 及 *Bra*023952 也有比较明显的响应（图 9）。

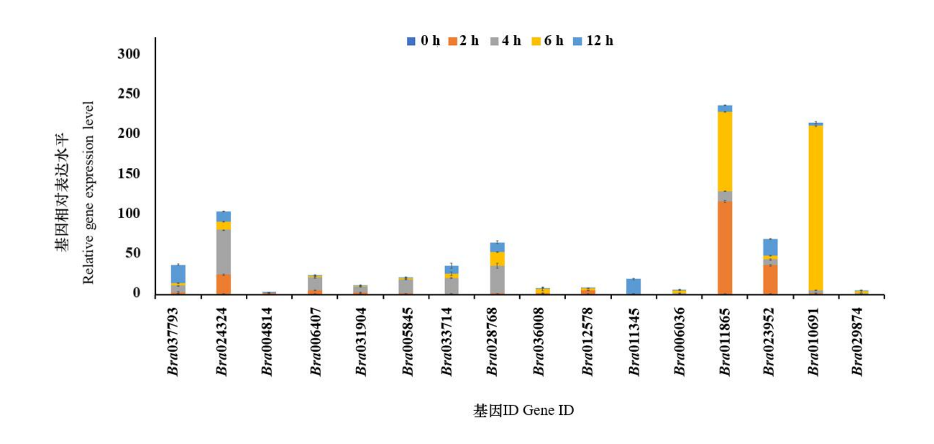


图 9 在盐胁迫下白菜 *CesA* 基因家族表达分析

Fig. 9 Expression analysis of *CesA* gene family under salt stress in *Brassica rapa*

3 讨论与结论

纤维素是植物细胞壁形成的主要成分，通过参与细胞形态建成，调控细胞发育，参与多种胞内信号转导途径，从而影响植物体的生长发育。植物纤维素合酶是由 36 个单体组成的，本质为糖基转移酶的玫瑰状复合体。Dobline et al.（2002）研究显示，纤维素合酶的主要功能是负责在质膜上催化纤维素的生物合成，高等植物纤维素合成酶复合体中，存在三种以上 CesA 亚型参与，其亚基之间相互作用形成高级复合物。本研究鉴定到 16 个白菜 *CesA* 基因，对白菜和拟南芥 *CesA* 基因家族共线性分析显示，两者的进化关系较近，而聚类在同一分支上的基因，如 *Bra*023952、*Bra*011345 与 *ATCesA*1（*AT*4*G*32410）聚类在同一分支上，可能具有相似的生物学功能。对白菜 *CesA* 基因家族成员进行的理化性质分析表明，蛋白等电点介于 4.76~9.12 之间，对酸性及碱性环境适应性均较好。16 个家族成员定位在 7 条染色体上，其中 ChrA03 上最多，且 *Bra*036008 定位在 scaffold 上。Nawaz et al. (2019) 的研究表明，家族成员蛋白有 8 个完整的跨膜结构域，各种植物来源的 *CesA* 基因在序列水平上有些差异。本研究中，白菜 *CesA* 家族成员多数有 8 个完整的跨膜结构域，其中 C 端有 6 个，近 N 端有 2 个，多数成员具有类似的保守结构域组成，表明其功能存在相似性。CesA 蛋白跨膜区在膜上形成一个通道用于分泌葡萄糖链，在分泌通道的活性区域的附近，DDD 残基与 QXXRW 保守结构域行使催化糖链合成的功能（Richmond, 2000）， 本研究显示，白菜 *CesA* 基因家族大部分成员均含有 DDD-QXXRW 结构域，与现有报道的拟南芥与烟草等具有相同的结构域，存在进化上的保守性。

CesA 在转录水平调控纤维素合成，编码序列中内含子的有无及其在序列中的位置造成了 *CesA* 基因间的差异（Richmond, 2000）。已有研究显示，烟草中鉴定出 21 个 *CesA* 基因家族成员中有 12~16 个内含子（Xu＆Kong, 2017），水稻中鉴定出 11 个家族成员中有 1~13 个内含子（Wang et al., 2015），铁皮石斛（Lan et al., 2019）和巨龙竹（Wang et al., 2021）中均含有较多内含子。本研究对基因结构的分析显示，白菜 *CesA* 基因家族成员都有 3~13 个内含子，因其内含子较多，基因进化保守性降低，转录时出现可剪接突变的概率较高，因此 *CesA* 基因家族之间在结构上存在差异，具有丰富的多样性。植物中存在多种纤维素合成酶基因，不同基因作用在不同部位，表达的功能有差异。南荻的 *MlCesA*7 参与初级细胞壁的形成，*MlCesA*4、*MlCesA*9 则参与次级细胞壁的形成（Li et al., 2021）;杉木的 *Cl CesA*1 参与木材次生壁形成，*ClCesA*2 参与细胞初生壁的形成（Pang et al., 2015）;琯溪蜜柚中，*CrCesA*2-2 基因参与初生壁纤维素合成调控，*CrCesA*4、*CrCesA*7-1、*CrCesA*8 参与了次生壁

纤维素的积累（Dai et al., 2021）;罗布麻的 *AvCesA*2、*AvCesA*3 参与初生细胞壁纤维素合成，*AvCesA*5、*AvCesA*8、*AvCesA*10 参与次生细胞壁纤维素合成（Xie et al., 2021）。在本研究中对二级结构的分析显示，除 *Bra*036008 和 *Bra*004814 外，白菜 *CesA* 基因家族成员均定位在质膜上，表明该基因家族与细胞壁纤维素合成相关。*CesA* 基因具有组织表达的特异性，Lan et al. (2019)对铁皮石斛 *CesA* 基因进行分析得出，*De*10015261、*De*10044519、*De*10126237 基因只在铁皮石斛的根中表达，*De*10015633、*De*10018596 只在茎中表达，*De*10002966 在根、叶、花中均有表达。本研究中，16 个白菜 *CesA* 家族基因在根茎叶花果荚中表达的个数和表达量均不相同，其中茎中表达量相对较高，尤其 *Bra*011865、*Bra*023952 和 *Bra*029874 在茎、叶、花中显著表达。Zhang et al. (2016) 的研究表明，*AtCesA*6 在拟南芥盐胁迫耐受性中发挥重要作用。本研究表明，大部分 *CesA* 基因对盐胁迫存在一定程度的响应，其中 *Bra*010691 的 6 h 处理组响应最为显著，表明该家族成员参与非生物胁迫的信号应答（Xu et al., 2017）。

目前对白菜 *CesA* 基因家族还没有进行系统的功能鉴定，许多基因的功能尚未发掘。本研究利用生物信息学方法对白菜 *CesA* 基因家族进行了全基因组鉴定，并对其进化关系、基因结构、顺式作用元件等进行分析，为进一步探究白菜生长发育机制奠定了基础。

参考文献

APPENZELLER L, DOBLIN M, BARREIRO R, et al., 2004. Cellulose synthesis in maize: isolation and express analysis of the cellulose synthase (*CesA*) gene family[J]. Cellulose, 11(3-4): 287-299.

BOLSER D, STAINES DM, PRITCHARD E, et al., 2003. Ensembl plants: Integrating tools for visualizing, mining, and analyzing plant genomics data[J]. Methods Mol Biol, 1374: 115-140.

Caño-Delgado A, PENFIELD S, SMITH C, et al., 2003. Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*[J]. Plant J, 34(3): 351-62.

CAI C, WANG X, LIU B, et al., 2017. Brassica rapa genome 2.0: A reference upgrade through sequence re-assembly and gene re-annotation. Mol Plant, 10(4):649-651.

CHEN C, CHEN H, ZHANG Y, et al., 2020. TBtools: An integrative toolkit developed for interactive analyses of big biological data[J]. Mol Plant, 13(8): 1194-1202.

DELMER DP, 1999. Cellulose biosynthesis: exciting times for a difficult field of study[J]. Annu Rev Plant Biol, 50: 245–276.

DOBLIN MS, KUREK I, JACOB-WILK D, et al., 2002. Cellulose biosynthesis in plants: from genes to rosettes[J]. Plant Cell Physiol, 43(12): 1407-20.

DAI YL, ZHAO QY, LIU RN, et al., 2021. Analysis of cellulose content and synthase gene expression in juice sacs secondary wall during granulation of Guanxi pomelo[J/OL]. Journal of Fruit Science: 1-13. [代亚兰，赵秋月，刘若南，等，2021. 琯溪蜜柚成熟期间汁胞纤维素含量及其合成酶[J/OL]. 果树学报：1-13.]

Hernández-Blanco C, FENG DX, HU J, et al., 2007. Impairment of cellulose synthases required for *Arabidopsis* secondary cell wall formation enhances disease resistance[J]. Plant Cell, 19(3): 890-903.

HUANG D, WANG S, ZHANG B, et al., 2015. A gibberellin-mediated DELLA-NAC signaling cascade regulates cellulose synthesis in rice[J]. Plant Cell, 27(6): 1681-96.

KUMAR S, Tamura K, Nei M, 1994. MEGA: Molecular evolutionary genetics analysis software for microcomputers[J]. Comput Appl Biosci, 10(2): 189-191.

KELLEY LA, MEZULIS S, YATES CM, et al., 2015. The Phyre2 web portal for protein modeling,

prediction and analysis[J]. Nat Protoc, 10(6): 845-858.
LI W, WANG J, SUN Q, et al., 2017. Expression analysis of genes encoding double B-box zinc finger proteins in maize[J]. Funct Integr Genomics, 17(6): 653-666.
LIU XQ, ZHANG H, GONG P, et al., 2018. Transcriptome analysis of secondary cell wall synthesis regulation at different developmental stages in Alfalfa (*Medicago sativa* L.)[J]. Scientia Agricultura Sinica, 51(11): 2049-2059. [刘希强，张涵，龚攀，等，2018. 紫花苜蓿不同发育时期次生壁合成调控的转录组分析[J]. 中国农业科学，51 (11): 2049-2059.]
LAN XT, HU S, FENG L, et al., 2019. Bioinformatic analysis of the cellulose synthase supergene family in *Dendrobium candidum*[J]. J Henan Agric Sci, 48(08): 49-55. [兰晓天，胡淞，冯磊，等，2019. 铁皮石斛纤维素合成酶超级基因家族生物信息学分析[J]. 河南农业科学，48(08): 49-55.]
LI Y, HE YQ, LIN YH, et al., 2021. Bioinformatics analysis of cellulose synthase CesA gene from *Miscanthus lutarioriparius*[J]. Mol Plant Breed, 19(13): 4378-4385. [李遥，何彦岑，林宇环，等，2021. 南荻纤维素合成酶 *CesA* 基因的生物信息学分析[J]. 分子植物育种, 19(13): 4378-4385.]
MENG CS, WANG ZW, ZHANG JH, et al., 2012. Bioinformatic comparison of the cellulose synthase gene family of cotton and *Arabidopsis thaliana*[J]. Guizhou Agric Sci, 40(07): 39-41. [孟成生，王志伟，张俊红，等，2012. 棉花与拟南芥纤维素合成酶基因家族的生物信息学比较[J]. 贵州农业科学，40(07): 39-41.]
NAWAZ, MALIN, XACHAN TF, et al., 2019. Characterization of cellulose synthase A (*CESA*) gene family in eudicots[J]. Biochem Genet, 57(2): 248-272.
PEAR JR, KAWaAGOE Y, SCHRECKENGOST WE, et al., 1996. Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. Proc Natl Acad Sci USA, 93(22): 12637-42.
PANG J, TONG ZK, HUANG HH, et al., 2015. Isolation and expression analysis of cellulose synthase genes in Chinese fir（*Cunninghamia lanceolata* [J]. J Zhejiang A & F Univ, 32(1): 40-46. [庞景，童再康，黄华宏，等，2015. 杉木纤维素合成酶基因 *CesA* 的克隆及表达分析[J]. 浙江农林大学学报，32(1): 40-46.]
RICHMOND T, 2000. Higher plant cellulose synthases[J]. Genome Biol, 1(4): 3001-3006.
TE RGL, XU K, LIN XF, BAI YE, 2021. Cloning of and bioinformatics analysis of cellulose synthase (*CesA*) genes from *Larix gmelinii* [J/OL]. Mol Breed: 1-18. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=FZZW20210330001&uniplatform=NZKPT&v=_MB1PS_Ch2DjioVqcoYL_Pw5MRDUtZjcgTYfFhYI1UjrI0jT5vN9tgs8CTJmX_3-. [特日格勒，徐坤，林晓飞，白玉娥，2021. 兴安落叶松纤维素合酶基因的克隆及生物信息学分析[J/OL]. 分子植物育种:1-18. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=FZZW20210330001&uniplatform=NZKPT&v=_MB1PS_Ch2DjioVqcoYL_Pw5MRDUtZjcgTYfFhYI1UjrI0jT5vN9tgs8CTJmX_3-.]
WONG HC, FEAR AL, CALHOON RD, 1990. Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. Proc Natl Acad Sci USA, 87(20): 8130-4.
WANG X, WANG H, WANG J, et al., 2011. *Brassica rapa* genome sequencing project consortium. The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet, 43(10):1035-9.
WANG Y, TANG H, DEBARRY JD, 2012. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity[J]. Nucleic Acids Res, 40(7): e49.

WANG ZY, WANG JP, PAN YX, et al., 2015. Bioinformatic analysis of *CesA* gene family in *Arabidopsis* and Rice[J]. J Henan Agric Sci, 44(06): 13-17. [王振怡，王金朋，潘玉欣，等，2015. 拟南芥和水稻 *CesA* 基因家族的生物信息学分析[J]. 河南农业科学，44(06): 13-17.]

WANG WZ, WANG J, WANG CM, et al, 2021. Bioinformatics analysis of cellulose synthase a (*CesA*) gene family based on the transcriptome of *Dendrocalamus sinicus* [J]. Mol Breed, 19(9): 2912-2921. [王文治，王娟，王昌命，等，2021. 基于转录组的巨龙竹 *CesA* 基因家族的生物信息学分析[J]. 分子植物育种，19(9): 2912-2921.]

XU ZC, KONG YZ, 2017. Genome-wide identification, subcellular localization and gene expression analysis of the members of CESA gene family in common tobacco (*Nicotiana tabacum* L.)[J]. Hereditas, 39(6): 512-524. [徐宗昌，孔英珍，2017. 普通烟草 *CESA* 基因家族成员的鉴定、亚细胞定位及表达分析[J]. 遗传，39(6): 512-524.]

XIE S, LI GQ, SONG LX, et al., 2021. Bioinformatics analysis of *Apocynum venetum CesA* gene family[J]. Guihaia, 41(4): 522-534. [解盛，李国旗，宋立肖，等，2021. 罗布麻 *CesA* 基因家族的生物信息学分析[J]. 广西植物，41(4): 522-534.]

YAO DY, WANG JZ, 1988. Cytoderm[J]. Chin Bull Bot, 1988(1): 18-21. [姚敦义，王静之，1988. 细胞壁[J]. 植物学通报，1988(1): 18-21.]

ZHANG SS, SUN L, DONG XR, et al., 2016. Cellulose synthesis genes *CESA6* and *CSI1* are important for salt stress tolerance in *Arabidopsis*[J]. J Integr Plant Biol, 58(7): 623-626.

ZHANG L, CAI X, WU J, et al.,2018. Improved *Brassica rapa* reference genome by single-molecule sequencing and chromosome conformation capture technologies. Hortic Res, 15;5:50.

ZHANG FL, YU SC, YU YJ, et al., 2021. Research progress on Chinese cabbage genetic breeding during 'The Thirteenth Five-year Plan' in China[J]. Chin Veget, 2021(1): 22-32. [张凤兰，于拴仓，余阳俊，等，2021. "十三五"我国大白菜遗传育种研究进展[J]. 中国蔬菜，2021(1): 22-32.]